REMARKS

Applicants appreciate the thorough examination of the present application as evidenced by the Office Action of December 23, 2008 (hereinafter "Office Action"). In response, Applicants have amended independent Claims 1, 4, and 13 as indicated above to clarify the recitations thereof. In particular, independent Claims 1 and 13 have been amended to clarify that each of the "means" recited therein is a "device," and that the noise reduction device or circuit recited therein is configured to separate a speaker's voice from the background noise "by removing said background noise from said audio sequence." Support for this amendment can be found, for example, at Page 8, lines 1-14 of the present specification. Also, independent Claim 4 has been amended to clarify that the voice of a speaker is detected from said discrete signal spectrum by analyzing visual features extracted from "a video sequence associated with extracted and analyzed audio features of the audio sequence," and that a noise power density spectrum of statistically distributed background noise is estimated "based on a signal that represents the voice of the speaker." Support for this amendment can be found, for example, at Figure 3b, steps S8a and S8b of the present specification. The dependent claims have also been amended for consistency with the amendments to independent Claims 1, 4, and 13. No new matter has been added.

Accordingly, Applicants respectfully request further consideration of the pending claims for at least the reasons discussed below.

Independent Claims 1 and 13 Are Patentable Over Basu and Girod

Claims 1-3 and 13-15 stand rejected under 35 U.S.C. 103(a) as being unpatentable over U.S. Patent Application Publication No. 2003/0018475 to Basu et al. ("Basu") and in view of U.S. Patent No. 6,483,532 to Girod ("Girod"). Amended Claim 1, for example, recites, in part:

> **a noise reduction circuit configured to separate a speaker's voice from said background noise** based on a combination of derived speech characteristics **by removing said separated background noise from said analog audio sequence** and configured to output a speech activity indication signal comprising a combination of speech activity estimates supplied by said audio feature extraction and analysis device and said visual feature extraction and analysis device; and
>
> a multi-channel acoustic echo cancellation unit configured to perform a near-end speaker detection and double-talk detection algorithm

> based on the speech characteristics derived by said audio feature extraction and analysis device and said visual feature extraction and analysis device. (*Emphasis added*).

Applicant respectfully submits that the combination of Basu and Girod does not disclose or suggest at least the above-highlighted recitations of pending Claim 1. Basu teaches an audio-visual speech detection and recognition system that provides speech recognition by discriminating between extraneous audible activity (such as background noise or background speech) that is not intended to be decoded, and speech that is intended to be decoded. *See* Basu, Paragraph 0094. In particular, as shown in Figure 8A, Basu uses visual information (e.g., mouth openings) to decide whether or not to decode an input audio signal. *See* Basu, Fig. 8A and Paragraph 0096. Thus, Basu describes discrimination between wanted speech and noise/unwanted speech to avoid "junk" recognition, for example, by enabling speech recognition when speech is detected, and disabling speech recognition when noise or unwanted speech is detected. *See* Basu, Paragraph 0096. However, such techniques for *distinguishing* speech from background noise in an input audio signal, as described in the cited portions of Basu, do not disclose or suggest *separating* speech from background noise in the input audio signal "by removing said background noise from said analog audio sequence," as recited by pending Claim 1.

Nor does the Office Action rely on Girod as disclosing or suggesting such a noise reduction circuit. *See* Office Action, Page 3. Furthermore, while Girod may disclose an echo cancellation circuit used in a video-assisted audio processing system (*see* Girod, Abstract), Applicants submit that it would not be obvious to use such an echo cancellation circuit in conjunction with the speech recognition techniques of Basu to "reliably distinguish" between desired speech and background speech, as asserted by the Office Action. *See* Office Action, Page 4. Indeed, as noted above, Basu uses detection of visual information (such as mouth movement) to determine whether a speech event to be decoded is occurring (*see* Basu, Paragraphs 0094 and 0096); as such, the echo cancellation circuit of Girod would not aid in such a determination based on visual information. Furthermore, Girod also uses detection of visible motion (such as mouth movement) to trigger the echo cancellation circuit described therein to filter the audio signal (*see* Girod, Col. 1, line 62 to Col. 2, line 7); thus, even if the echo cancellation circuit of Girod were included in the system of Basu, such a combination would not avoid false activation of the system, as the system would be activated prior to (or

simultaneously with) performing the echo cancellation responsive to the detected mouth movement. Thus, Applicants submit that it would not be obvious to selectively combine techniques for distinguishing speech from noise in a given audio signal based on visual information, as described in Basu, with techniques for cancelling echo from the audio signal, as described in Girod.

Accordingly, Applicant submits that pending Claim 1 is patentable over the combination of Basu and Girod for at least the above reasons. Pending Claim 13 includes similar recitations, and is thus patentable for at least similar reasons. Dependent Claims 2-3 and 14-15 are also patentable at least per the patentability of Claims 1 and 13 from which they depend.

Independent Claim 4 Is Patentable Over Basu and Wynn

Claims 4, 7 and 8 stand rejected under 35 U.S.C. 103(a) as being unpatentable over Basu and in view of U.S. Patent No. 5,706,394 to Wynn (hereinafter "Wynn").

> 4. A near-end speaker detection method for reducing noise in a detected analog audio sequence, said method comprising:
> converting said analog audio sequence into a digital audio sequence;
> calculating a corresponding discrete signal spectrum of the digital audio sequence by performing a Fast Fourier Transform (FFT);
> detecting a voice of a speaker from said discrete signal spectrum by analyzing visual features extracted from a video sequence associated with extracted and analyzed audio features of the audio sequence, the visual features including current locations of face, lip movements and/or facial expressions of the speaker in a sequence of images in the video sequence;
> **estimating a noise power density spectrum** of statistically distributed background noise **based on a signal that represents the voice of the speaker**;
> subtracting a discretized version of the estimated noise power density spectrum from the discrete signal spectrum of the digital audio sequence to obtain a difference signal; and
> calculating a corresponding discrete time-domain signal of the obtained difference signal by performing an Inverse Fast Fourier Transform (IFFT) to provide a recognized speech signal. (*Emphasis added*).

As recited by pending Claim 4, in some embodiments, a noise power density spectrum of the statistically distributed background noise may be estimated based on a signal

that represents the speaker's voice, e.g., as illustrated, for example, in Fig. 3b, step S8a of the present specification. *See* Specification, Fig. 3b. Such an estimation of a noise power density spectrum based on the voice of the speaker may provide improved and/or more reliable noise estimation, especially where the speaker's voice is detected based on analyzing visual features extracted from a video sequence associated with the audio sequence.

Applicant respectfully submits that the combination of Basu and Wynn does not disclose or suggest at least the above-highlighted recitations of pending Claim 1. Wynn describes a signal processing system for filtering noise using a voice activity detector (VAD) to detect frames containing voice/speech and noise-only frames. *See* Wynn, Abstract and Col. 8, line 65-Col. 9, line 11. However, Wynn notes that the noise spectrum described therein "is estimated from noise-only frames detected by VAD 25." Wynn, Col. 4, lines 53-55. Wynn further notes that "the noise-only frames detected between speech segments are used to update the noise power spectrum estimate," and that "[e]stimating the noise power spectral density $S_d(f)$ from noise-only frames using a voice activity detector (VAD)...is based on the assumption that the noise present during speech has the same average power spectrum as the estimated $S_d(f)$." Wynn, Col. 9, lines 23-39. Accordingly, while Wynn may describe operations for estimating a noise power spectral density, nowhere does Wynn disclose or suggest that the noise power spectral density described therein is "based on a signal that represents the voice of the speaker," as recited by pending Claim 4. Moreover, as Wynn specifically describes estimating the noise spectrum based on noise-only frames, rather than frames containing the voice of the speaker, Applicants submit that Wynn teaches away from the recitations of pending Claim 4.

Nor does the Office Action rely on Basu as disclosing or suggesting such noise estimation. *See* Office Action, Pages 5-6. Accordingly, Applicant submits that pending Claim 4 is patentable over the combination of Basu and Wynn for at least the above reasons. Dependent Claims 5-12 are also patentable at least per the patentability of Claim 4 from which they depend.

The Information Disclosure Statement

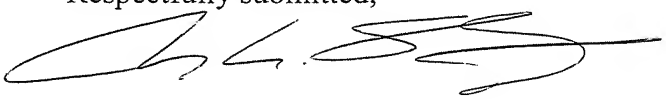
The Office Action asserts that Applicant's Information Disclosure Statement (IDS) of July 20, 2005 fails to comply with 37 CFR 1.98(a)(2), due to a failure to provide copies of the items 5 and 6 (the International Search Report ("ISR") and the International Preliminary

Examination Report ("IPER") for the corresponding European Application, respectively) cited therein. *See* Office Action, Page 8. However, Applicant respectfully note that item 5 (the ISR) and item 6 (the IPER) of the IDS of July 20, 2005 currently appear in PAIR, with a mail room date of July 20, 2005 (*see* 10th and 11th items from the bottom, entitled "Documents submitted with 371 Applications"). Accordingly, consideration of these documents is respectfully requested.

Conclusion

Accordingly, based on the above amendments and remarks, Applicant submits that the pending claims are now in condition for allowance. Thus, Applicant respectfully requests allowance of these claims and passing the application to issue. Applicant encourages the Examiner to contact the undersigned to resolve any remaining issues.

Respectfully submitted,



Rohan G. Sabapathypillai
Registration No.: 51,074

USPTO Customer No. 54414
Myers Bigel Sibley & Sajovec
Post Office Box 37428
Raleigh, North Carolina 27627
Telephone: 919/854-1400
Facsimile: 919/854-1401

CERTIFICATION OF ELECTRONIC TRANSMISSION UNDER 37 CFR § 1.8

I hereby certify that this correspondence is being transmitted via the Office electronic filing system in accordance with § 1.6(a)(4) to the U.S. Patent and Trademark Office on March 20, 2009.

Audra Wooten